

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: Fruskin, Michael J.  
Fingler, Douglas J.  
Licarella, Dominic  
Newman, Walter
- (ii) TITLE OF INVENTION: Mucosal Vascular Addressins and Uses Thereof
- (iii) NUMBER OF SEQUENCES: 17
- (iv) CORRESPONDENCE ADDRESS:  
(A) ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
(B) STREET: Two Elm Street Drive  
(C) CITY: Lexington  
(D) STATE: Massachusetts  
(E) COUNTRY: U.S.A.  
(F) TEL: 617-211-4779
- (v) COMPUTER READABLE FORM:  
(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: Patent In Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:  
(A) APPLICATION NUMBER: US 08/875,849  
(B) FILING DATE: 11-FEB-1996  
(C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:  
(A) APPLICATION NUMBER: US 08/523,004  
(B) FILING DATE: 11-SEP-1995
- (viii) PRIOR APPLICATION DATA:  
(A) APPLICATION NUMBER: US 08/386,357  
(B) FILING DATE: 10-FEB-1995
- (viii) ATTORNEY/AGENT INFORMATION:  
(A) NAME: Prosser, David E.  
(B) REGISTRATION NUMBER: 21,592  
(C) REFERENCE, DOCKET NUMBER: LKS94-04A2
- (ix) TELECOMMUNICATION INFORMATION:  
(A) TELEPHONE: 781-861-8240  
(B) TELEFAX: 781-861-9540

(2) INFORMATION FOR SEQ ID NO:1:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1624 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(11) MOLECULE TYPE: cDNA

(12) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1218

(21) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATG GAT TTC GGA CTC GGC CTC CTG CTG GCG GGG CTT CTG GGG CTC CTC 48  
Met Asp Phe Gly Leu Ala Leu Leu Leu Ala Gly Leu Leu Gly Leu Leu  
1 10 15

CTC GGC CAG TCC CTC CAG GTG AAG CCC CTG CAG GTG GAG CCC CCG GAG 96  
Leu Gly Gln Ser Leu Gln Val Lys Pro Leu Gln Val Glu Pro Pro Glu  
20 25 30

CCG GTG GTG GAA GTG GGC TTG GGC GCG TCG CGC CAG CTC ACC TGC CGC 144  
Pro Val Val Ala Val Ala Leu Gly Ala Ser Arg Gln Leu Thr Cys Arg  
35 40 45

CTG GGC TGC GAA GAC CGC GGC GGC TCC GTG CAG TGG CGG GGC CTC GAC 192  
Leu Ala Cys Ala Asp Arg Gly Ala Ser Val Gln Trp Arg Gly Leu Asp  
50 55 60

ACC AGC CTG GGA GCA GTC CAG TCG GAC ACG GGC CGC AGC GTC CTC ACC 240  
Thr Ser Leu Gly Ala Val Gln Ser Asp Thr Gly Arg Ser Val Leu Thr  
65 70 75 80

GTG CGC AAC GGC TCG CTG TCG GGC GGC ACC CGC GTG TGC GTG GGC 288  
Val Arg Asn Ala Ser Leu Ser Ala Ala Gly Thr Arg Val Cys Val Gly  
85 90 95

TCC TGC GGC GGC CGC ACC TTC CAG CAC ACC GTC CAG CTC CTT GTG TAC 336  
Ser Cys Gly Gly Arg Thr Phe Gln His Thr Val Gln Leu Leu Val Tyr  
100 105 110

GCC TTC CCG GAC CAG CTG ACC GTC TCC CCA GCA ACC CTC GTG CCG GGT 384  
Ala Phe Pro Asp Gln Leu Thr Val Ser Pro Ala Ala Leu Val Pro Gly  
115 120 125

GAC CGC GAG GTG GGC TGT ACC GGC CAC AAA GTC ACG CCC GTG GAC CCC 432  
Asp Pro Glu Val Ala Cys Thr Ala His Lys Val Thr Pro Val Asp Pro  
130 135 140

AAC GCG CTC TCC TTC TTC CTG CTC GTC GCG GCG CAG GAA CTG GAG GCG 480  
Asn Ala Leu Ser Phe Ser Leu Leu Val Gly Gly Gln Glu Leu Glu Gly  
145 150 155 160

GGG CAA GGC CTG GGC CCG GAG GTG CAG GAG GAG GAG GAG GAG GGC CAG	528
Ala Gln Ala Leu Gly Pro Glu Val Gln Glu Glu Glu Glu Glu Pro Gln	
165 170 175	
GGG GAG GAG GAG GTG CTG TTC AGG GTG ACA GAG CCG TGG CCG CTG CCG	576
Gly Asp Glu Asp Val Leu Phe Arg Val Thr Gln Arg Trp Arg Leu Pro	
180 185 190	
CCC CTG GGG ACC CCT GTC CCG CCC GGC CTC TAC TGC CAG GGC ACG ATG	624
Pro Leu Gly Thr Pro Val Pro Pro Ala Leu Tyr Cys Gln Ala Thr Met	
195 200 205	
AGG CTG CCT GGC TTG CAG CTC AGC CAC CGC CAG GGC ATC CCC GTC CTG	672
Arg Leu Pro Gly Leu Glu Leu Ser His Arg Gln Ala Ile Pro Val Leu	
210 215 220	
CAC AGC CCG ACC TCC CCG GAG CCT CCC GAC ACC ACC TCC CCG GAG CCT	720
His Ser Pro Thr Ser Pro Glu Pro Pro Asp Thr Thr Ser Pro Glu Pro	
225 230 235 240	
CCC AAC ACC ACC TCC CCG GAG TCT CCC GAC ACC ACC TCC CCG GAG TCT	768
Pro Asn Thr Thr Ser Pro Glu Ser Pro Asp Thr Thr Ser Pro Glu Ser	
245 250 255	
CCC GAC ACC ACC TCC CAG GAG CCT CCC GAC ACC ACC TCC CAG GAG CCT	816
Pro Asp Thr Thr Ser Gln Glu Pro Pro Asp Thr Thr Ser Gln Glu Pro	
260 265 270	
CCC GAC ACC ACC TCC CAG GAG CCT CCC GAC ACC ACC TCC CCG GAG CCT	864
Pro Asp Thr Thr Ser Gln Glu Pro Pro Asp Thr Thr Ser Pro Glu Pro	
275 280 285	
CCC GAC AAG ACC TCC CCG GAG CTC GGC CCC CAG CAG GGC TCC ACA CAG	912
Pro Asp Lys Thr Ser Pro Glu Pro Ala Pro Gln Gln Gly Ser Thr His	
290 295 300	
ACC CCC AAG ACC TCC CCG GAG CTC GGC CCC CAG CAG GGC TCC ACA CAG	960
Thr Pro Arg Ser Pro Gly Ser Thr Arg Thr Arg Arg Pro Glu Ile Ser	
305 310 315 320	
CAG GCT GGG CCC ACG CAG GGA GAA GTG ATC CCA ACA GGC TCC TCC AAA	1008
Gln Ala Gly Pro Thr Gln Gly Glu Val Ile Pro Thr Gly Ser Ser Lys	
325 330 335	
CCT GCG GGT GAC CAG CTG CCG GCG GGT CTG TGG ACC AGC AGT GCG GTG	1056
Pro Ala Gly Asp Gln Leu Pro Ala Ala Leu Trp Thr Ser Ser Ala Val	
340 345 350	
CTG GGA CTG CTG CTC CTG GGC TTG CCC ACG TAT CAC CTC TGG AAA CGC	1104
Leu Gly Leu Leu Leu Leu Ala Leu Pro Thr Tyr His Leu Trp Lys Arg	
355 360 365	
TGC CGG CAC CTG GCT GAG GAC GAC ACC CAC CCA CCA GCT TCT CTG AGG	1152
Cys Arg His Leu Ala Glu Asp Asp Thr His Pro Pro Ala Ser Leu Arg	
370 375 380	

mb  
D  
nt

B

CTT CTG CCC CAG GTG TCG GCC TGG GCT GGG TTA AGG GGG ACC GGC CAG 1200  
 Leu Leu Pro Gln Val Ser Ala Trp Ala Gly Leu Arg Gly Thr Gly Gln  
 390 395 400  
 GTC GGG ATC AGC CCC TCC TGAGTGGCCA GCCTTCCGC CTGTGAAAG 1248  
 Val Gly Ile Ser Pro Ser  
 405  
 AAAATAGCTT GGACCCCTTC AAGTTGAGAA CTGGTCAGGG CAAACCTGCC TCCATTCTA 1308  
 CTCAAAGTCA TCCCTCTGCT CACAGAGATG GATGCATGTT CTGATTGCCT CTTTGGAGAA 1368  
 GCTCATCAGA AACTCAAAAG AAGGCCACTG TTTGTCTCAC CTACCCATGA CTTGAAGGCC 1428  
 CTCCCTGAGT GGTCGCCACC TTCTGGACG GAACCACGTA CTTTTTACAT ACATTGATTC 1488  
 ATGTCTCAGG TCTCCCTAAA AATGCCAAG ACCAAGCTGT GCCCTGACCA CCGTGGGCC 1548  
 CTGTCTCAG GACCTCCTGA GCCTTGGCA AATAAACCTC CTAAATGAT AAAAAAAAAA 1608  
 AAAAAAAAAA AA/AAA 1624

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
- A) LENGTH: 406 amino acids
  - B) TYPE: amino acid
  - C) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Asp Phe Gly Leu Ala Leu Leu Leu Ala Gly Leu Leu Gly Leu Leu 15  
 1 5 10  
 Leu Gly Gln Ser Leu Gln Val Lys Pro Leu Gln Val Glu Pro Pro Glu 30  
 20 25 30  
 Pro Val Val Ala Val Ala Leu Gly Ala Ser Arg Gln Leu Thr Cys Arg 45  
 35 40 45  
 Leu Ala Cys Ala Asp Arg Gly Ala Ser Val Gln Trp Arg Gly Leu Asp 60  
 50 55 60  
 Thr Ser Leu Gly Ala Val Gln Ser Asp Thr Gly Arg Ser Val Leu Thr 80  
 65 70 75 80  
 Val Arg Asn Ala Ser Leu Ser Ala Ala Gly Thr Arg Val Cys Val Gly 95  
 85 90 95  
 Ser Cys Gly Gly Arg Thr Phe Gln His Thr Val Gln Leu Leu Val Tyr 110  
 100 105 110  
 Ala Phe Pro Asp Gln Leu Thr Val Ser Pro Ala Ala Leu Val Pro Gly 125  
 115 120 125

Asp Pro Glu Val Ala Cys Thr Ala His Lys Val Thr Pro Val Asp Pro  
130 135 140

Asn Ala Leu Ser Phe Ser Leu Leu Val Gly Gly Gln Glu Leu Glu Gly  
145 150 155 160

Ala Glu Ala Leu Gly Pro Glu Val Gln Glu Glu Glu Glu Glu Pro Gln  
165 170 175

Gly Asp Glu Asp Val Leu Phe Arg Val Thr Glu Arg Trp Arg Leu Pro  
180 185 190

Pro Leu Gly Thr Pro Val Pro Pro Ala Leu Tyr Cys Gln Ala Thr Met  
195 200 205

Arg Leu Pro Gly Leu Glu Leu Ser His Arg Gln Ala Ile Pro Val Leu  
210 215 220

His Ser Pro Thr Ser Pro Glu Pro Pro Asp Thr Thr Ser Pro Glu Pro  
225 230 235 240

Pro Asn Thr Thr Ser Pro Glu Ser Pro Asp Thr Thr Ser Pro Glu Ser  
245 250 255

Pro Asp Thr Thr Ser Gln Glu Pro Pro Asp Thr Thr Ser Gln Glu Pro  
260 265 270

Pro Asp Thr Thr Ser Gln Glu Pro Pro Asp Thr Thr Ser Pro Glu Pro  
275 280 285

Pro Asp Lys Thr Ser Pro Glu Pro Ala Pro Gln Gln Gly Ser Thr His  
290 295 300

Thr Pro Arg Ser Pro Gly Ser Thr Arg Thr Arg Arg Pro Glu Ile Ser  
305 310 315 320

Gln Ala Gly Pro Thr Gln Gly Glu Val Ile Pro Thr Gly Ser Ser Lys  
325 330 335

Pro Ala Gly Asp Gln Leu Pro Ala Ala Leu Trp Thr Ser Ser Ala Val  
340 345 350

Leu Gly Leu Leu Leu Leu Ala Leu Pro Thr Tyr His Leu Trp Lys Arg  
355 360 365

Cys Arg His Leu Ala Glu Asp Asp Thr His Pro Pro Ala Ser Leu Arg  
370 375 380

Leu Leu Pro Gln Val Ser Ala Trp Ala Gly Leu Arg Gly Thr Gly Gln  
385 390 395 400

Val Gly Ile Ser Pro Ser  
405

Pro  
D11

B

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1539 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) FEATURES:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1146

(iv) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATG GAT TTC GGA CTG GCG CTC CTG CTG GCG GGG CTT CTG GGG CTC CTC 48  
Met Asp Phe Gly Leu Ala Leu Leu Leu Ala Gly Leu Leu Gly Leu Leu  
1 5 10 15

CTC GGC CAG TCC CTC CAG GTG AAG CCC CTG CAG GTG GAG CCC CCG GAG 96  
Leu Gly Gln Ser Leu Gln Val Lys Pro Leu Gln Val Glu Pro Pro Glu  
2 25 30

CCG GTG GTG GCG GTG GCG TTG GGC GCG TCC CGC CAG CTC ACC TGC CGC 144  
Pro Val Val Ala Val Ala Leu Gly Ala Ser Arg Gln Leu Thr Cys Arg  
35 40 45

CTG GCG TGC GCG GAC CGC GCG GCG TCG GTG CAG TGG CGG GGC CTG GAC 192  
Leu Ala Cys Ala Asp Arg Gly Ala Ser Val Gln Trp Arg Gly Leu Asp  
50 55 60

ACC AGC CTC GGC GCG GTG CAG TCG GAC ACC GCG CGC AGC GTC CTC ACC 240  
Thr Ser Leu Gly Ala Val Gln Ser Asp Thr Gly Arg Ser Val Leu Thr  
65 70 75 80

GTG CGC AAC GCG TCG CTG TCG GCG GCG GGG ACC CGC GTG TGC GTG GGC 288  
Val Arg Asn Ala Ser Leu Ser Ala Ala Gly Thr Arg Val Cys Val Gly  
85 90 95

TCC TGG GGG GCG CGC ACC TTC CAG CAG ACC GTG CAG CTC CTT GTG TAC 336  
Ser Cys Gly Gly Arg Thr Phe Gln His Thr Val Gln Leu Leu Val Tyr  
100 105 110

GCC TTC CCG GAC CAG CTG ACC GTC TCC CTA GCA GCG CTG GTG CCT GGT 384  
Ala Phe Pro Asp Gln Leu Thr Val Ser Pro Ala Ala Leu Val Pro Gly  
115 120 125

GAC CCG GAG GTG GCC TGT ACG GCC CAG AAA GTC ACG CCC GTG GAC CCC 432  
Asp Pro Glu Val Ala Cys Thr Ala His Lys Val Thr Pro Val Asp Pro  
130 135 140

AAC GCG CTC TCC TTC TCC CTG CTC GTC GGG GCG CAG GAA CTG GAG GGG 480  
Asn Ala Leu Ser Phe Ser Leu Leu Val Gly Gly Gln Glu Leu Glu Gly  
145 150 155 160

GGG CAA GTC CTG GGC CCG GAG GTG CAG GAG GAG GAG GAG GAG GAG GAG  
Ala Gln Ala Leu Gly Pro Glu Val Gln Glu Glu Glu Glu Glu Glu Glu  
165 170 175

GGG GAG GAG GAG GAG GTG CTG TTC AGG GTG ACA GAG GAG GAG GAG GAG GAG GAG  
Gly Asp Glu Asp Val Leu Phe Arg Val Thr Glu Ala Trp Arg Leu Pro  
180 185 190

GGG CTG GGC ACC CCT GTC CCG CCC GGC CTG TAC TGC CAG GGC AGC ATG  
Pro Leu Gly Thr Pro Val Pro Pro Ala Leu Tyr Cys Gln Ala Thr Met  
195 200 205

AGG CTG CCT GGC TTG GAG CTC AGC CAC CCG CAG GGC ATC CCC GTC CTG  
Arg Leu Pro Gly Leu Glu Leu Ser His Arg Gln Ala Ile Pro Val Leu  
210 215 220

CAC AGC CCG ACC TCC CCG GAG CCT CCC GAC ACC ACC TCC CCG CAG TCT  
His Ser Pro Thr Ser Pro Glu Pro Pro Asp Thr Thr Ser Pro Glu Ser  
225 230 235 240

CCC GAC ACC ACC TCC CCG GAG TCT CCC GAC ACC ACC TCC CAG GAG CCT  
Pro Asp Thr Thr Ser Pro Glu Ser Pro Asp Thr Thr Ser Gln Glu Pro  
245 250 255

CCC GAC ACC ACC TCC CCG GAG CCT CCC GAC AAG ACC TCC CCG GAG CCC  
Pro Asp Thr Thr Ser Pro Glu Pro Pro Asp Lys Thr Ser Pro Glu Pro  
260 265 270

GCC CCC CAG CAG GGC TCC ACA CAG ACC CCC AGG AGC CCA GGC TCC ACC  
Ala Pro Gln Gln Gly Ser Thr His Thr Pro Arg Ser Pro Gly Ser Thr  
275 280 285

AGG ACT CGC GGC CCT GAG ATC TCC CAG GGT GGC CCC AGC CAG GGA GAA  
Arg Thr Arg Arg Pro Glu Ile Ser Gln Ala Gly Pro Thr Gln Gly Glu  
290 295 300

GTG ATC CCA ACA GGC TCG TCC AAA CCT GGC GGT GAG CAG CTG CCC GCG  
Val Ile Pro Thr Gly Ser Ser Lys Pro Ala Gly Asp Gln Leu Pro Ala  
305 310 315 320

GCT CTG TGG ACC AGC AGT GCG GTG CTG GGA CTG CTG TCC CTG GCG TTG  
Ala Leu Trp Thr Ser Ser Ala Val Leu Gly Leu Leu Leu Leu Ala Leu  
325 330 335

CCC ACC TAT CAC CTC TGG AAA CCC TCC CCG CAG CTG CCT GAG GAG GAG  
Pro Thr Tyr His Leu Trp Lys Arg Cys Arg His Leu Ala Glu Asp Asp  
340 345 350

ACC CAC CCA CCA GGT TCT CTG AGC CTT CTG CCC CAG CTG TCG GCC TGG  
Thr His Pro Pro Ala Ser Leu Arg Leu Leu Pro Gln Val Ser Ala Trp  
355 360 365

GCT GGG TTA AGC GGC ACC GGC CAG GTC GGC ATC AGC CCC TCC  
Ala Gly Leu Arg Gly Thr Gly Gln Val Gly Ile Ser Pro Ser  
370 375 380

TGAGTGGCCA GACTTTCCG CTGTGAAAGC AAAATAGCTT GACCCCTTC AAGTTGAGAA  
1206

CTGGTCAGGG CAAACCTCC TCCATTCTA CTCAAAGTCA TCCCTCTGTT CACAGAGATG  
1266

GATGCATGTT CTGATTGCCT CTTTGGAGAA GCTCATCAGA AACTCAAAAG AAGGCCATTC 1326  
 TTTTCTCAAT CTACCATCA CTGAAGCCC CTCCCTGACT GGTCCCCAGC TTTCTGACG 1386  
 CAACACGTA CTTTTACAT ACATTGATTC ATGTCTCAGG TCTCCCTAAA AATGCTAAG 1446  
 ACAAGCTGT GGCCTACCA CCGTGGGTC CTCTGCTCA GACCTCTTA GCTTTCCTTA 1506  
 AATAACCTC CTAAATGAA AAAAAAAAAA AAA 1566

(2) INFORMATION FOR SEQ ID NO:4:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 382 amino acids  
 (B) TYPE: amino acid  
 (C) TOPOLOGY: linear

(2) MOLECULE TYPE: protein

(3) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Asp Phe Gly Leu Ala Leu Leu Leu Ala Gly Leu Leu Gly Leu Leu  
 1 5 10 15  
 Leu Gly Gln Ser Leu Gln Val Lys Pro Leu Gln Val Glu Pro Pro Glu  
 20 25 30  
 Pro Val Val Ala Val Ala Leu Gly Ala Ser Arg Gln Leu Thr Cys Arg  
 35 40 45  
 Leu Ala Cys Ala Asp Arg Gly Ala Ser Val Gln Trp Arg Gly Leu Asp  
 50 55 60  
 Thr Ser Leu Gly Ala Val Gln Ser Asp Thr Gly Arg Ser Val Leu Thr  
 65 70 75 80  
 Val Arg Asn Ala Ser Leu Ser Ala Ala Gly Thr Arg Val Cys Val Gly  
 85 90 95  
 Ser Cys Gly Gly Arg Thr Phe Gln His Thr Val Gln Leu Leu Val Tyr  
 100 105 110  
 Ala Phe Pro Asp Gln Leu Thr Val Ser Pro Ala Ala Leu Val Pro Gly  
 115 120 125  
 Asp Pro Glu Val Ala Cys Thr Ala His Lys Val Thr Pro Val Asp Pro  
 130 135 140  
 Asn Ala Leu Ser Phe Ser Leu Leu Val Gly Gly Gln Glu Leu Glu Gly  
 145 150 155 160  
 Ala Gln Ala Leu Gly Pro Glu Val Gln Glu Glu Glu Glu Pro Gln  
 165 170 175  
 Gly Asp Glu Asp Val Leu Phe Arg Val Thr Glu Arg Trp Arg Leu Pro  
 180 185 190



Pro Leu Gly Thr Pro Val Pro Pro Ala Leu Tyr Cys Gln Ala Thr Met  
195 200 205

Arg Leu Pro Gly Leu Glu Leu Ser His Arg Gln Ala Ile Pro Val Leu  
210 215 220

His Ser Pro Thr Ser Pro Glu Pro Pro Asp Thr Thr Ser Pro Gln Ser  
225 230 235 240

Pro Asp Thr Thr Ser Pro Glu Ser Pro Asp Thr Thr Ser Gln Glu Pro  
245 250 255

Pro Asp Thr Thr Ser Pro Glu Pro Pro Asp Lys Thr Ser Pro Glu Pro  
260 265 270

Ala Pro Gln Gln Gly Ser Thr His Thr Pro Arg Ser Pro Gly Ser Thr  
275 280 285

Arg Thr Arg Arg Pro Glu Ile Ser Gln Ala Gly Pro Thr Gln Gly Glu  
290 295 300

Val Ile Pro Thr Gly Ser Ser Lys Pro Ala Gly Asp Gln Leu Pro Ala  
305 310 315 320

Ala Leu Trp Thr Ser Ser Ala Val Leu Gly Leu Leu Leu Leu Ala Leu  
325 330 335

Pro Thr Tyr His Leu Trp Lys Arg Cys Arg His Leu Ala Glu Asp Asp  
340 345 350

Thr His Pro Pro Ala Ser Leu Arg Leu Leu Pro Gln Val Ser Ala Trp  
355 360 365

Ala Gly Leu Arg Gly Thr Gly Gln Val Gly Ile Ser Pro Ser  
370 375 380

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1721 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

- (ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 4..1038

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AGC ATG GAT CGG GGC CTG ACC CTC CTG CTG GCG GGG CTT CTG GGG CTC  
 Met Asp Arg Gly Leu Ala Leu Leu Leu Ala Gly Leu Leu Gly Leu  
 1 5 10 15

CTC CAG CCC GGC TGC GGC CAG TCC CTC CAG GTG AAG CCC CTG CAG GTG 96  
 Leu Gln Pro Gly Cys Gly Gln Ser Leu Gln Val Lys Pro Leu Gln Val  
 20 25 30

GAG CCC CCC GAG CCC GTG GTG GCC GTG GGC CTG GGC GGC TCT CAC CAG 144  
 Gln Pro Pro Glu Pro Val Val Ala Val Ala Leu Gly Ala Ser Ala Gln  
 35 40 45

CTC ACC TGC CGC CTG GAC TCC CCG GAC CCC GGC GCC ACC GTG CAG TCG 192  
 Leu Thr Cys Arg Leu Asp Cys Ala Asp Arg Gly Ala Thr Val Gln Thr  
 50 55 60

CGG GGC CTC GAC ACC AGC CTG GGC GCG GTG CAG TCG GAC GCG GGC CGC 240  
 Arg Gly Leu Asp Thr Ser Leu Gly Ala Val Gln Ser Asp Ala Gly Arg  
 65 70 75

AGC GTC CTC ACC GTG CGC AAC GCC TCG CTG TCG GCG GCC GCG ACC CGT 288  
 Ser Val Leu Thr Val Arg Asn Ala Ser Leu Ser Ala Ala Gly Thr Arg  
 80 85 90 95

GTG TGC GTG GGC TCC TGC GGC GGC CGC ACC TTC CAG CAC ACC GTG CGG 336  
 Val Cys Val Gly Ser Cys Gly Gly Arg Thr Phe Gln His Thr Val Arg  
 100 105 110

CTC CTT GTG TAC GCT TTC CCG GAC CAG CTG ACC ATC TCC CCG GCA GCC 384  
 Leu Leu Val Tyr Ala Phe Pro Asp Gln Leu Thr Ile Ser Pro Ala Ala  
 115 120 125

CTG GTG CCT GGT GAC CCG GAG GTG GGC TGT ACG GGC CAC AAA GTC ACG 432  
 Leu Val Pro Gly Asp Pro Ser Val Ala Cys Thr Ala His Lys Val Thr  
 130 135 140

CCT GTG GAC CCG AAT GGC CTC TCC TTC TTC CTG CTT CTG GGG GAC CAG 480  
 Pro Val Asp Pro Asn Ala Leu Ser Phe Ser Leu Leu Leu Gly Asp Gln  
 145 150 155

GAA CTG GAG GGC GCG CAG GGT CTG GGC CTG GAG GTG GAG GAG GAG GAG 528  
 Glu Leu Gln Gly Ala Gln Ala Leu Gly Pro Gly Val Gln Glu Glu Glu  
 160 165 170 175

GAG GAG CCC CAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG 576  
 Glu Glu Pro Gln Glu Glu Glu Asp Val Leu Phe Arg Val Thr Glu Arg  
 180 185 190

TGG GAG CTG CCG ACC CTG GCA AAT CCT GTG CTG CCG GCG CTC TAC TGC 624  
 Trp Arg Leu Pro Thr Leu Ala Thr Pro Val Leu Pro Ala Leu Tyr Cys  
 195 200 205

CAG GGC ACC ATG AGC CTG CCT GGC TTC GAG CTC AAT CAG CCG CAG GCC 672  
 Gln Ala Thr Met Arg Leu Pro Gly Leu Gln Leu Ser His Arg Gln Ala  
 210 215 220

ATC CCG GTG CTG CAG GGC CCG ACC TTC GGC GAG CCC CCC GAG ACG ACC 720  
 Ile Pro Val Leu His Gly Pro Thr Ser Arg Glu Pro Pro Asp Thr Thr  
 225 230 235

TCC CCG GAA CCC CGC GCG GCG ACC TTC CCG GAG ACC ACC CCC CAG CAG 768  
 Ser Pro Glu Pro Arg Ala Ala Thr Ser Pro Glu Thr Thr Pro Gln Gln  
 240 245 250 255

GGC TCC ACA CAC AGC CCC AGG AGC CCG GCC TCT ACC AGG ACT TGC CCG 816  
 Gly Ser Thr His Ser Pro Arg Ser Pro Gly Ser Thr Arg Thr Cys Arg  
 260 265 270  
 CCT GAG ATC TCC CAG GCT GCG CCC AGG CAG GGA GAA GTG AAT CTA ATA 864  
 Pro Glu Ile Ser Gln Ala Gly Pro Thr Gln Gly Glu Val Ile Pro Thr  
 275 280 285  
 GGC TCC TCC AAA CCT ACG GCT GAC CAG CTG CCC GCG GCT CTG TCG ACC 912  
 Gly Ser Ser Lys Pro Thr Gly Asp Gln Leu Pro Ala Ala Leu Trp Thr  
 290 295 300  
 AGC AGT GCG GTG CTG CGA CTG CTG CTC CTG GCT TTG CCC ACC TAC CAC 960  
 Ser Ser Ala Val Leu Gly Leu Leu Leu Leu Ala Leu Pro Thr Tyr His  
 305 310 315  
 CTC TGG AAA CGT TGC CGG CAC CTG GCT GAG GAC GGC GCC CAC CCA CCA 1008  
 Leu Trp Lys Arg Cys Arg His Leu Ala Glu Asp Gly Ala His Pro Pro  
 320 325 330 335  
 GCT TCT CTG AGT AGC CAG CCC TCC CCC CTG TGAAGGGAAA ATAGGTTGGA 1058  
 Ala Ser Leu Ser Ser Gln Pro Phe Pro Leu  
 340 345  
 CCCCTTCAAG CTGAGAGCTG GTGCGGGGAAA ACCTGCCTCC CATTCTATTC AAAGTCATCG 1118  
 CTCTGGTCAC AGAGAGGAC GCACATTCTG ATTGCCTCCT TTGGAAAGGC TCATCAGAAA 1178  
 CTCAAAAGAA GTTATCTTT TGTTCGCTT ACCGTGACC TGGAAAGCCCC CGCCCCGCTC 1238  
 GAGTGACCCC TGAATTCTG GAGGAAACA AGGTACTTCT TACATATATT GATTTCATGTG 1298  
 TCATATCTCC CTAATAATGCG TAAATCAGC TGTGCTCCGA CCACCTTGGG CCCCTGCCAT 1358  
 CAGGAACCTCC TGAGGCTTG GAAATAAAG CTCTAAAAG GATAGAACT GAAACTTGTG 1418  
 GCGGGGCGCG GTGGCTAAG CTTGTAATCC CAGCACTTTG GGAGGCGAG GTGGTGTGAT 1478  
 CACGAGSTCA GGAGATGAG ACCATCTCTG CTAACCCGTG AAACCCCGTG TCTACTAAAA 1538  
 AAATACAAAA ATTAGCGGG AGCGGTGCG GGCGCTGTA GTCCAGCTA CTCGGAGGC 1598  
 TGAGGCAGGA GAATGGCTG AACCGGAGG GCGGAGCTG CAGTGAGCTG AGATCCGSCC 1658  
 ACTGCACTCC AGCTGCGGG ACAGAGCGAG ACTCCCTCTC AAAAAAAAAA AAAAAAAAAA 1718  
 AAA 1721

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 345 amino acids
  - (E) TYPE: amino acid
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: protein

(X1) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Asp Arg Gly Leu Ala Leu Leu Leu Ala Gly Leu Leu Gly Leu Leu  
1 5 10

Gln Pro Gly Cys Gly Gln Ser Leu Gln Val Lys Pro Leu Gln Val Gln  
20 25 30

Pro Pro Gln Pro Val Val Ala Val Ala Leu Gly Ala Ser Arg Gln Leu  
35 40 45

Thr Cys Arg Leu Asp Cys Ala Asp Arg Gly Ala Thr Val Gln Trp Arg  
50 55 60

Gly Leu Asp Thr Ser Leu Gly Ala Val Gln Ser Asp Ala Gly Arg Ser  
65 70 75 80

Val Leu Thr Val Arg Asn Ala Ser Leu Ser Ala Ala Gly Thr Arg Val  
85 90 95

Cys Val Gly Ser Cys Gly Gly Arg Thr Phe Gln His Thr Val Arg Leu  
100 105 110

Leu Val Tyr Ala Phe Pro Asp Gln Leu Thr Ile Ser Pro Ala Ala Leu  
115 120 125

Val Pro Gly Asp Pro Glu Val Ala Cys Thr Ala His Lys Val Thr Pro  
130 135 140

Val Asp Pro Asn Ala Leu Ser Phe Ser Leu Leu Leu Gly Asp Gln Glu  
145 150 155 160

Leu Glu Gly Ala Gln Ala Leu Gly Pro Glu Val Glu Glu Glu Glu Glu  
165 170 175

Glu Pro Gln Glu Glu Glu Asp Val Leu Phe Arg Val Thr Glu Arg Trp  
180 185 190

Arg Leu Pro Thr Leu Ala Thr Pro Val Leu Pro Ala Leu Tyr Cys Gln  
195 200 205

Ala Thr Met Arg Leu Pro Gly Leu Glu Leu Ser His Arg Gln Ala Ile  
210 215 220

Pro Val Leu His Gly Pro Thr Ser Arg Glu Pro Pro Asp Thr Thr Ser  
225 230 235 240

Pro Glu Pro Arg Ala Ala Thr Ser Pro Glu Thr Thr Pro Gln Gln Gly  
245 250 255

Ser Thr His Ser Pro Arg Ser Pro Gly Ser Thr Arg Thr Cys Arg Pro  
260 265 270

Glu Ile Ser Gln Ala Gly Pro Thr Gln Gly Glu Val Ile Pro Thr Gly  
275 280 285

Ser Ser Lys Pro Thr Gly Asp Gln Leu Pro Ala Ala Leu Trp Thr Ser  
290 295 300

Ser Ala Val Leu Gly Leu Leu Leu Leu Ala Leu Pro Thr Tyr His Leu  
305 310 315 320

Trp Lys Ala Cys Arg His Leu Ala Glu Asp Gly Ala His Pro Pro Ala  
325 330 335

Ser Leu Ser Ser Glu Pro Phe Pro Leu  
340 345

(2) INFORMATION FOR SEQ ID NO:7:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: unknown

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CTCTACTGCC AGGGAAG

18

(2) INFORMATION FOR SEQ ID NO:8:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: unknown

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AGCCTGGGA3 ATCTAAGG

19

(2) INFORMATION FOR SEQ ID NO:9:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: unknown

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GCCACGATGA AGCTGCTTG

20

(2) INFORMATION FOR SEQ ID NO:10:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: unknown

(81) SEQUENCE DESCRIPTION: SEQ ID NO:10:

1998, 1999, 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023, 2024, 2025, 2026, 2027, 2028, 2029, 2030, 2031, 2032, 2033, 2034, 2035, 2036, 2037, 2038, 2039, 2040, 2041, 2042, 2043, 2044, 2045, 2046, 2047, 2048, 2049, 2050, 2051, 2052, 2053, 2054, 2055, 2056, 2057, 2058, 2059, 2060, 2061, 2062, 2063, 2064, 2065, 2066, 2067, 2068, 2069, 2070, 2071, 2072, 2073, 2074, 2075, 2076, 2077, 2078, 2079, 2080, 2081, 2082, 2083, 2084, 2085, 2086, 2087, 2088, 2089, 2090, 2091, 2092, 2093, 2094, 2095, 2096, 2097, 2098, 2099, 2100, 2101, 2102, 2103, 2104, 2105, 2106, 2107, 2108, 2109, 2110, 2111, 2112, 2113, 2114, 2115, 2116, 2117, 2118, 2119, 2120, 2121, 2122, 2123, 2124, 2125, 2126, 2127, 2128, 2129, 2130, 2131, 2132, 2133, 2134, 2135, 2136, 2137, 2138, 2139, 2140, 2141, 2142, 2143, 2144, 2145, 2146, 2147, 2148, 2149, 2150, 2151, 2152, 2153, 2154, 2155, 2156, 2157, 2158, 2159, 2160, 2161, 2162, 2163, 2164, 2165, 2166, 2167, 2168, 2169, 2170, 2171, 2172, 2173, 2174, 2175, 2176, 2177, 2178, 2179, 2180, 2181, 2182, 2183, 2184, 2185, 2186, 2187, 2188, 2189, 2190, 2191, 2192, 2193, 2194, 2195, 2196, 2197, 2198, 2199, 2200, 2201, 2202, 2203, 2204, 2205, 2206, 2207, 2208, 2209, 2210, 2211, 2212, 2213, 2214, 2215, 2216, 2217, 2218, 2219, 2220, 2221, 2222, 2223, 2224, 2225, 2226, 2227, 2228, 2229, 2230, 2231, 2232, 2233, 2234, 2235, 2236, 2237, 2238, 2239, 2240, 2241, 2242, 2243, 2244, 2245, 2246, 2247, 2248, 2249, 2250, 2251, 2252, 2253, 2254, 2255, 2256, 2257, 2258, 2259, 2260, 2261, 2262, 2263, 2264, 2265, 2266, 2267, 2268, 2269, 2270, 2271, 2272, 2273, 2274, 2275, 2276, 2277, 2278, 2279, 2280, 2281, 2282, 2283, 2284, 2285, 2286, 2287, 2288, 2289, 2290, 2291, 2292, 2293, 2294, 2295, 2296, 2297, 2298, 2299, 2300, 2301, 2302, 2303, 2304, 2305, 2306, 2307, 2308, 2309, 2310, 2311, 2312, 2313, 2314, 2315, 2316, 2317, 2318, 2319, 2320, 2321, 2322, 2323, 2324, 2325, 2326, 2327, 2328, 2329, 2330, 2331, 2332, 2333, 2334, 2335, 2336, 2337, 2338, 2339, 2340, 2341, 2342, 2343, 2344, 2345, 2346, 2347, 2348, 2349, 2350, 2351, 2352, 2353, 2354, 2355, 2356, 2357, 2358, 2359, 2360, 2361, 2362, 2363, 2364, 2365, 2366, 2367, 2368, 2369, 2370, 2371, 2372, 2373, 2374, 2375, 2376, 2377, 2378, 2379, 2380, 2381, 2382, 2383, 2384, 2385, 2386, 2387, 2388, 2389, 2390, 2391, 2392, 2393, 2394, 2395, 2396, 2397, 2398, 2399, 2400, 2401, 2402, 2403, 2404, 2405, 2406, 2407, 2408, 2409, 2410, 2411, 2412, 2413, 2414, 2415, 2416, 2417, 2418, 2419, 2420, 2421, 2422, 2423, 2424, 2425, 2426, 2427, 2428, 2429, 2430, 2431, 2432, 2433, 2434, 2435, 2436, 2437, 2438, 2439, 2440, 2441, 2442, 2443, 2444, 2445, 2446, 2447, 2448, 2449, 2450, 2451, 2452, 2453, 2454, 2455, 2456, 2457, 2458, 2459, 2460, 2461, 2462, 2463, 2464, 2465, 2466, 2467, 2468, 2469, 2470, 2471, 2472, 2473, 2474, 2475, 2476, 2477, 2478, 2479, 2480, 2481, 2482, 2483, 2484, 2485, 2486, 2487, 2488, 2489, 2490, 2491, 2492, 2493, 2494, 2495, 2496, 2497, 2498, 2499, 2500, 2501, 2502, 2503, 2504, 2505, 2506, 2507, 2508, 2509, 2510, 2511, 2512, 2513, 2514, 2515, 2516, 2517, 2518, 2519, 2520, 2521, 2522, 2523, 2524, 2525, 2526, 2527, 2528, 2529, 2530, 2531, 2532, 2533, 2534, 2535, 2536, 2537, 2538, 2539, 2540, 2541, 2542, 2543, 2544, 2545, 2546, 2547, 2548, 2549, 2550, 2551, 2552, 2553, 2554, 2555, 2556, 2557, 2558, 2559, 2560, 2561, 2562, 2563, 2564, 2565, 2566, 2567, 2568, 2569, 2570, 2571, 2572, 2573, 2574, 2575, 2576, 2577, 2578, 2579, 2580, 2581, 2582, 2583, 2584, 2585, 2586, 2587, 2588, 2589, 2590, 2591, 2592, 2593, 2594, 2595, 2596, 2597, 2598, 2599, 2600, 2601, 2602, 2603, 2604, 2605, 2606, 2607, 2608, 2609, 2610, 2611, 2612, 2613, 2614, 2615, 2616, 2617, 2618, 2619, 2620, 2621, 2622, 2623, 2624, 2625, 2626, 2627, 2628, 2629, 2630, 2631, 2632, 2633, 2634, 2635, 2636, 2637, 2638, 2639, 2640, 2641, 2642, 2643, 2644, 2645, 2646, 2647, 2648, 2649, 2650, 2651, 2652, 2653, 2654, 2655, 2656, 2657, 2658, 2659, 2660, 2661, 2662, 2663, 2664, 2665, 2666, 2667, 2668, 2669, 2670, 2671, 2672, 2673, 2674, 2675, 2676, 2677, 2678, 2679, 26

(2) INFORMATION FOR SEQ ID NO:11:

(1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 32 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGAAGCTTCC ACCA.TGGT.TT TCGGAC TGGC CC

(2) INFORMATION FILE SER. ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 16 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CCGACTAGTG TCGGACCTTG TAGGAC

(2) INFORMATION FILE SER. ID NO:13:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 17 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xix) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GGACTAGTGG TTAGAG GAG CCGTGTG

(2) INFORMATION FOR SEQ ID NO:14:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (C) UNFAMDELINNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Gly Leu Asp Thr Ser Leu  
1 5

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 6 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ix) FEATURE:  
A: NAME, KEY: Region  
B: LOCATION: 2  
D: OTHER INFORMATION: /product= "OTHER"  
/label= variable  
/note= "Xaa = Ile or Leu"

(ix) FEATURE:  
A: NAME, KEY: Region  
B: LOCATION: 3  
D: OTHER INFORMATION: /product= "OTHER"  
/label= variable  
/note= "Xaa = Asp or Glu"

(ix) FEATURE:  
A: NAME, KEY: Region  
B: LOCATION: 4  
D: OTHER INFORMATION: /product= "OTHER"  
/label= variable  
/note= "Xaa = Thr or Ser"

(ix) FEATURE:  
A: NAME, KEY: Region  
B: LOCATION: 5  
D: OTHER INFORMATION: /product= "OTHER"  
/label= variable  
/note= "Xaa = Pro or Ser"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Gly Xaa Xaa Xaa Xaa Leu  
1 5

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 6 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Gln Ile Asp Ser Pro Leu  
1 5

(ii) INFORMATION FOR SEQ ID NO:17:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 5 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Leu Asp Thr Ser Leu  
1 5